BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structural and biophysical characterization of an epitope-specific engineered Fab fragment and complexation with membrane proteins: implications for co-crystallization

Jennifer L. Johnson,[a] Kevin C. Entzminger,[b] Jeongmin Hyun,[b] Sibel Kalyoncu,[a] David P. Heaner Jr,[a] Ivan A. Morales,[a] Aly Sheppard,[a] James C. Gumbart,[a,c] Jennifer A. Maynard[b]* and Raquel L. Lieberman[a]*

[a]School of Chemistry and Biochemistry, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, GA 30332, USA, [b]McKetta Department of Chemical Engineering, University of Texas at Austin, MC0400, 1 University Station, Austin, TX 78712, USA, and [c]School of Physics, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, GA 30332, USA. *Correspondence e-mail: maynard@che.utexas.edu, raquel.lieberman@chemistry.gatech.edu

Received 22 November 2014
Accepted 28 January 2015

Edited by J. L. Martin, University of Queensland, Australia

Keywords: Fab fragment; crystallization chaperones; membrane proteins; GPCRs; β-barrel.

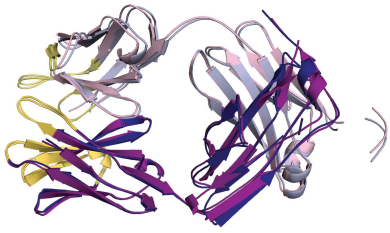
PDB reference: engineered Fab fragment, 4x0k

Supporting information: this article has supporting information at journals.iucr.org/d

Crystallization chaperones are attracting increasing interest as a route to crystal growth and structure elucidation of difficult targets such as membrane proteins. While strategies to date have typically employed protein-specific chaperones, a peptide-specific chaperone to crystallize multiple cognate peptide epitope-containing client proteins is envisioned. This would eliminate the target-specific chaperone-production step and streamline the co-crystallization process. Previously, protein engineering and directed evolution were used to generate a single-chain variable (scFv) antibody fragment with affinity for the peptide sequence EYMPME (scFv/EE). This report details the conversion of scFv/EE to an anti-EE Fab format (Fab/EE) followed by its biophysical characterization. The addition of constant chains increased the overall stability and had a negligible impact on the antigen affinity. The 2.0 Å resolution crystal structure of Fab/EE reveals contacts with larger surface areas than those of scFv/EE. Surface plasmon resonance, an enzyme-linked immunosorbent assay, and size-exclusion chromatography were used to assess Fab/EE binding to EE-tagged soluble and membrane test proteins: namely, the β-barrel outer membrane protein intimin and α-helical A2a G protein-coupled receptor (A_2aR). Molecular-dynamics simulation of the intimin constructs with and without Fab/EE provides insight into the energetic complexities of the co-crystallization approach.

1. Introduction

Crystallization of a given target protein can be a difficult step in structure determination and is especially notorious in the case of membrane proteins (MPs). Numerous strategies have emerged to enhance the crystallizability of difficult proteins. These include protein modifications such as random mutagenesis, directed evolution or strategic mutations to improve crystallizability (Pédelacq *et al.*, 2002; Yang *et al.*, 2003; Keenan *et al.*, 2005; Argos *et al.*, 1979), generation or discovery of ligands to stabilize a specific protein conformation (Vedadi *et al.*, 2006; Connelly, 1994), protein symmetrization by cross-linking or engineered metal-binding sites (Laganowsky *et al.*, 2011; Banatao *et al.*, 2006), limited proteolysis to remove flexible regions (Wernimont & Edwards, 2009) and surface-entropy reduction (Derewenda, 2004), among others. Recent improvements specific to MPs include the use of lipidic cubic phase (LCP) and lipid-mimicking detergents to stabilize MPs in a native-like environment to promote crystal growth (Caffrey, 2003; Privé, 2009; Howell *et al.*, 2010)



For membrane proteins, the so-called crystallization chaperone approach has been particularly successful. The desirable biophysical properties of chaperones include an increase in hydrophilic residues available for forming crystal contacts, thus improving the likelihood of obtaining well ordered crystals of the chaperone–target membrane protein complex. Covalent chaperones have been utilized to crystallize several G protein-coupled receptors (GPCRs), in which intracellular loop 3 (ICL3; Cherezov *et al.*, 2007; Wu *et al.*, 2010; Liu *et al.*, 2012) or the N-terminus (Zou *et al.*, 2012; Thompson *et al.*, 2012) is replaced by T4 lysozyme (T4L) or apocytochrome b_{562} RIL (BRIL) and generated as a chimeric protein. Noncovalent chaperones include antibody fragments, mostly Fab and scFv generated by hybridoma or library screening (Kovari *et al.*, 1995; Ostermeier *et al.*, 1995; Hunte *et al.*, 2000; Zhou *et al.*, 2001; Stura *et al.*, 2002; Rasmussen *et al.*, 2007; Tereshko *et al.*, 2008; Uysal *et al.*, 2009), V_{HH} camelid domains (nanobodies; Rasmussen *et al.*, 2011; Hassaine *et al.*, 2014) and ankyrin-repeat proteins (DARPins; Huber *et al.*, 2007; Sennhauser & Grütter, 2008). All of these noncovalent binding partners are target-specific; namely, for each new protein of interest a new chaperone must be sought.

Previously, we proposed the use of peptide-specific antibody fragments as an alternative to protein-specific crystallization chaperones (Lieberman *et al.*, 2011). To the best of our knowledge, the only successful example of a target-independent non-covalent chaperone involved the use of a Fab fragment that recognizes a portable small structural RNA element to crystallize a ribozyme (Koldobskaya *et al.*, 2011; Piccirilli & Koldobskaya, 2011). Co-crystallization with client membrane proteins has been attempted by us using engineered scFvs (Pai *et al.*, 2011; Kalyoncu *et al.*, 2014) and by others using Fab fragments generated from commercially available FLAG-binding monoclonal antibodies (Roosild *et al.*, 2006), but without documented success to date. Analysis of deposited co-crystal structures of membrane proteins using crystallization chaperones in the Protein Data Bank (PDB) reveals that the majority rely on the Fab format (Lieberman *et al.*, 2011). This observation could be due to the increased surface area available to Fab fragments to form crystal contacts and/or the increased stability that the constant domains provide (Röthlisberger *et al.*, 2005) compared with scFvs.

Here, we report the generation of a Fab fragment with a nanomolar affinity for the EE epitope (Fab/EE) along with detailed structural and biochemical characterization relevant to its potential use as a crystallization chaperone. We report successful complex formation with a variety of proteins containing EE peptides, including two EE-tagged maltose-binding protein constructs, the α-helical membrane protein human adenosine A2a G protein-coupled receptor (A_2aR-GFP-EE) and the *Escherichia coli* β-barrel membrane protein intimin (intimin-EE1 and intimin-EE2). Molecular-dynamics simulations with wild-type (WT) intimin, intimin-EE1, intimin-EE2 and the Fab/EE–intimin-EE complexes reveal an unexpected increase in flexibility when mutating native loop residues to the EE epitope, which is likely to be hindering the crystallization of the complex. The implications of these findings for co-crystallization are discussed.

2. Materials and methods

2.1. Molecular biology, expression and purification of Fab/EE

To convert our previously described 3D5/EE_48 scFv (scFv/EE; Pai *et al.*, 2011) to the Fab format, the variable light and heavy chains were sequentially subcloned *via* polymerase chain reaction (PCR) into the NcoI–NotI and NheI–HindIII restriction sites, respectively, of the pFab vector (courtesy of Dr Georgiou, University of Texas at Austin, USA; Levy *et al.*, 2001), resulting in the vector pFab-Fab/EE. The vector provides in-frame, N-terminal periplasmic leader sequences and C-terminal peptide tags: a decahistidine tag on the light chain and a FLAG tag on the heavy chain (Supplementary Tables S1 and S2). The fidelity of the construct was confirmed by DNA sequencing (University of Texas Austin core facility) using the primers listed in Supplementary Table S3. Fab/EE was expressed in *E. coli* BL21 (DE3) cells. 2 ml Luria–Bertani broth (LB; Fisher) culture supplemented with 60 µg ml^{-1} ampicillin was inoculated with a single colony and incubated for ∼4 h at 37°C with shaking at 225 rev min^{-1}. The starter culture was diluted 1:100 in 500 ml Terrific Broth (TB; Fisher) in a 2 l baffled flask and grown overnight with shaking at 225 rev min^{-1} and 25°C. Cells were pelleted at 4200*g* for 10 min and 4°C, and the cell pellet was then resuspended in 500 ml fresh TB in a 2 l flask and incubated for 1 h at 25°C and 225 rev min^{-1} before inducing expression with 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG; Calbiochem) for 4.5–5 h. The cells were pelleted and flash-frozen in liquid nitrogen before placing them in a −80°C freezer or were subjected to lysis directly. Fab/EE was purified as reported previously for scFv/EE (Maynard *et al.*, 2005). Briefly, the cell pellet was resuspended in 10 ml resuspension buffer (0.1 *M* Tris pH 8.0, 0.75 *M* sucrose) per gram of cell pellet. Osmotic shock was carried out by adding 7.5 ml 1 m*M* EDTA and 2.5 mg lysozyme per gram of cells, rocking or stirring for 45 min to 1 h at 4°C, then adding 1 ml 0.5 *M* $MgCl_2$ per gram of cells and stirring for an additional 45 min to 1 h. After centrifugation for 20 min at 47 800*g* (SS-34 rotor), the supernatant was subjected to Ni^{2+}-affinity purification with a wash buffer consisting of 20 m*M* Tris pH 8, 500 m*M* NaCl, 20 m*M* imidazole and an elution buffer containing either 100 m*M* EDTA or 500 m*M* imidazole. Fab/EE was further purified by preparative size-exclusion chromatography (SEC) using a Superdex 75 16/600 column equilibrated with 50 m*M* HEPES pH 7.5, 150 m*M* NaCl (HBS; Supplementary Fig. S1) on an ÄKTA FPLC system (GE Healthcare).

2.2. Biophysical characterization of Fab/EE

For all proteins, protein purity and size were assessed by standard reducing and nonreducing 12% SDS–PAGE analysis (Sambrook & Russell, 2001) using silver stain for A_2aR and Coomassie stain for the visualization of all other proteins (Supplementary Fig. S1, inset), with protein concentrations

Table 1
Fab/EE data and refinement statistics.

Values in parentheses are for the outer shell.

Data collection	
Beamline and source	22-ID, APS
X-ray wavelength (Å)	1.0
Resolution (Å)	32.97–2.04 (2.11–2.04)
Space group	*P*1
Unit-cell parameters (Å, °)	a = 53.559, b = 67.131, c = 71.877, α = 71.3, β = 78.1, γ = 85.31
Total No. of reflections	226544
No. of unique reflections	57696 (5221)
Multiplicity	3.9 (3.9)
Completeness (%)	97.71 (93.40)
$\langle I/\sigma(I)\rangle$	14.66 (2.70)
$R_{\rm merge}$ (%)	8.4
Refinement statistics	
Final $R_{\rm cryst}$	0.1641 (0.2198)
Final $R_{\rm free}$†	0.2078 (0.2680)
No. of non-H atoms	
Protein	7012
Water	525
Total	7537
R.m.s. deviations	
Bonds (Å)	0.004
Angles (°)	0.9
Average B factors (Å^2)	
Overall	39.1
Protein	38.8
Water	44.3
Ramachandran plot‡	
Most favored (%)	97.7
Allowed (%)	2.2
Outliers (%)	0.1

† $R_{\rm free}$ is calculated for a randomly chosen 5% of reflections which were not used for structure refinement. ‡ As calculated by *RAMPAGE* (Lovell *et al.*, 2003).

determined by Micro BCA assay (Pierce) or estimated from the absorbance at 280 nm combined with calculated extinction coefficients based on amino-acid composition using *ProtParam* (Gasteiger *et al.*, 2005). Fractions of monomeric Fab/EE were pooled for subsequent experiments. Fab/EE solubility was determined as described previously (Pai *et al.*, 2011) by measuring the concentration of soluble protein remaining after concentration to ∼20 mg ml^{-1} and 4 d incubation at 4°C. Fab/EE thermal stability was measured by thermal unfolding after mixing 20 µl of 200 µ*M* Fab/EE or HBS-only control with SYPRO Orange (1 µl at 1:1000 dilution; Molecular Probes) in a real-time PCR machine (ViiA 7; Applied Biosystems) in increments of 0.96°C min^{-1} from 25 to 90°C. Analysis to determine the melting temperature ($T_{\rm m}$), the midpoint of unfolding, was performed with the *ViiA* 7 (Applied Biosystems) software.

2.3. Protein crystallization, data collection, structure determination and refinement

Fab/EE (6.5 mg ml^{-1} in HBS) was crystallized at room temperature by the sitting-drop vapour-diffusion method. Conditions were optimized based on Wizard I and II (Emerald Bio) solution G4 consisting of 20% polyethylene glycol (PEG) 8000, 100 m*M* MES pH 6.0, 200 m*M* calcium acetate. Crystals used for structure determination were grown from a reservoir solution consisting of 0.1 *M* HEPES pH 7.5, 100 m*M* calcium acetate, 20–26%(*w*/*v*) PEG 8000, 3% 1-propanol. The crystals were harvested and cryocooled in reservoir solution supplemented with 15% glycerol. Crystallographic data were collected on beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (APS), Illinois, USA. The data were indexed, integrated and scaled in *HKL*-3000 (Minor *et al.*, 2006). The structure was solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using a polyalanine search model prepared in *CHAINSAW* (Stein, 2008) derived from the Fab portion of PDB entry 3sob (Bourhis *et al.*, 2011). The model was iteratively rebuilt in *Coot* (Emsley *et al.*, 2010) and refined in *PHENIX* (Adams *et al.*, 2010). Ramachandran outliers were determined using *RAMPAGE* (Lovell *et al.*, 2003). Crystallographic statistics are presented in Table 1 and the structure was deposited in the PDB as entry 4x0k.

2.4. Computational analysis of Fab/EE crystal contacts

PDBe Protein Interfaces, Surfaces and Assemblies (*PDBePISA*; Krissinel & Henrick, 2007) was used to rank and analyze crystal lattice contacts by surface area and energy, as well as to catalog critical amino acids in crystal contacts based on their formation of hydrogen-bond or salt-bridge interactions. After excluding the native heavy–light chain interface within the Fab/EE monomer, the top three interfaces were identified as major crystal contacts and used in further analysis.

2.5. Molecular biology, expression and purification of proteins presenting the EE epitope

2.5.1. Soluble test proteins. All soluble test proteins used have been described previously (Pai *et al.*, 2011). The EE tag was appended to the C-terminus of maltose-binding protein (MBP-EE) or into a surface loop of MBP, replacing residues 170–175 (MBP-KEE). MBP with only a C-terminal hexahistidine tag was used as a negative control (MBP-His$_6$). An scFv with the EE tag inserted into the flexible linker region (scFv-EE$_1$) was used as another EE-tagged test protein for BIAcore experiments (see §2.6.1). These proteins were expressed and purified *via* the C-terminal hexahistidine tag as described for Fab/EE.

2.5.2. A$_2$aR expression and purification. The plasmid pITy-A$_2$aR-GFP-His$_{10}$ was generously provided by Dr Anne Robinson (University of Tulane). An EE-tagged variant, pITy-A$_2$aR-GFP-EE, was generated by insertion of the EE epitope after the wild-type residue Lys209 (numbering as in GenBank AAA83270) in the third intracellular loop (ICL3), flanked by GS residues to allow peptide accessibility (Supplementary Fig. S2). This was accomplished by site-directed mutagenesis (SDM) using the QuikChange mutagenesis kit (Stratagene). Colony PCR was used to screen for modified plasmids as described previously (Costa & Weiner, 2006). Correct EE-peptide insertion of the final plasmid was confirmed by DNA sequencing. All primers used are listed in Supplementary Table S3.

Plasmids pITy-A_2aR-GFP and pITy-A_2aR-GFP-EE were transformed into *Saccharomyces cerevisiae* BJ5464 by electroporation, with individual yeast colonies screened for high expression by whole-cell GFP fluorescence, and the highest-expressing clones were used for subsequent protein purification as described previously (O'Malley *et al.*, 2007). Briefly, cell lysis was accomplished by vortexing and the protein was purified by Ni^{2+} affinity chromatography using wash buffer consisting of 50 m*M* NaH_2PO_4, 300 m*M* NaCl, 10% glycerol, 1 m*M* phenylmethylsulfonyl fluoride (PMSF), 0.1% *n*-dodecyl-β-D-maltoside (DDM), 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 0.02% cholesteryl hemisuccinate (CHS) pH 8.0 with protease inhibitors, and the same buffer was used for elution with the addition of 500 m*M* imidazole. Ni^{2+}-affinity purification was followed by SEC using a Superdex 200 16/60 (GE Healthcare) column equilibrated with GPCR buffer (10 m*M* NaH_2PO_4 pH 7.0 with 0.1% DDM, 0.1% CHAPS, 0.02% CHS). The purified protein was analyzed by SDS–PAGE and Western blot using the primary anti-EE peptide Glu-Glu monoclonal antibody (1:1000 dilution in blocking buffer; Covance) and incubation with horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody (ThermoFisher). The signal was developed with SuperSignal West Dura Extended Duration Substrate (Thermo Scientific) and the resulting image was captured on X-ray film.

2.5.3. Intimin molecular biology, expression and purification. The plasmid for *E. coli* intimin (Fairman *et al.*, 2012) was generously provided by Dr Susan Buchanan (NIH). The EE epitope was incorporated into an extramembraneous loop in wild-type intimin as described previously (intimin-EE1; Kalyoncu *et al.*, 2014). Two alanines were inserted on either side of the EE tag *via* SDM using the QuikChange II mutagenesis kit (Agilent) and the primers listed in Supplementary Table S3 (intimin-EE2). Intimin-EE1 and intimin-EE2 were expressed and purified as previously described for wild-type intimin (Fairman *et al.*, 2012).

2.6. Binding assays and complex formation

2.6.1. Surface plasmon resonance (SPR). Kinetic binding assays were performed with a BIAcore 3000 (GE Healthcare) instrument using immobilized bovine serum albumin (BSA) or Fab/EE coupled to CM5 chips *via* NHS–EDC chemistry to a level of ~1200 RU as bait for ligand proteins. Responses owing to sample refractive-index changes and nonspecific binding were corrected using the signal from a flow cell coupled with BSA. Purified MBP-KEE, MBP-EE, scFv-EE_1 or control MBP-His_6 were injected in a duplicate dilution series from 2 to 0.125 μM at a flow rate of 50 $\mu l\ min^{-1}$ to minimize mass-transport effects in HBS running buffer supplemented with 0.005% Tween-20. Surface regeneration was performed after each run with a single 30 s injection of 2 *M* $MgCl_2$. GPCR variants were injected similarly with the exceptions of using a single dilution series from the highest concentration available upon purification and using a running buffer composed of HBS with 0.1% DDM. The association rate constant (k_{on}), dissociation rate constant (k_{off}) and equilibrium dissociation constant (K_d; $K_d = k_{off}/k_{on}$) were calculated assuming a Langmuir 1:1 binding model with the *BIAevaluation* software. Only data sets with $\chi^2 < 0.5$ were used. *GraphPad Prism* 5 was used for graphical representation.

2.6.2. Enzyme-linked immunosorbent assay (ELISA). Fab/EE binding to A_2aR-GFP-EE was also assessed by ELISA. High-binding 96-well plates (Costar) were coated with 20 $\mu g\ ml^{-1}$ Fab/EE overnight at 4°C. After a 1 h incubation with blocking buffer at room temperature, purified A_2aR-GFP-EE or A_2aR-GFP proteins were serially diluted twofold in blocking buffer (PBS with 5% milk). This was followed by 1 h incubation with rabbit anti-GFP (Invitrogen), washing, and a 1.5 h incubation with goat anti-rabbit HRP (Sigma). After a final wash step, TMB substrate (Vector Labs) was added, the signal was allowed to develop and the reaction was quenched with 1 *M* H_2SO_4. Absorbance was read at 450 nm on a SpectraMax M5 microplate reader (Molecular Devices). Data points represent the average of at least two measurements, including error bars equal to one standard deviation. *GraphPad Prism* 5 was used to fit data to a three-parameter logistic model and for graphical representation.

2.6.3. Size-exclusion chromatography (SEC). Fab/EE–client protein interactions were further evaluated by SEC fractionation on an ÄKTA FPLC system (GE Healthcare) at 4°C. Equimolar amounts of Fab/EE and MBP-KEE were incubated for 90 min either together or separately at room temperature prior to fractionation on a Superdex S200 column (GE Healthcare) equilibrated with HBS. For complexation with intimin, Fab/EE in 50 m*M* Tris pH 7.5, 200 m*M* NaCl, 0.01% sodium azide was combined with WT intimin, intimin-EE1 or intimin-EE2 in a 1:1 molar ratio and incubated for 2 h on ice before injection onto a Superose 12 10/300 GL column equilibrated with 50 m*M* Tris pH 7.5, 200 m*M* NaCl, 0.01% sodium azide, 0.05% DDM. Elution fractions for each peak were either concentrated by Amicon Ultra (30K MWCO, Millipore) or precipitated by tricholoracetic acid (TCA; Link & LaBaer, 2011) and analyzed by 12% reducing SDS–PAGE.

2.7. Molecular-dynamics (MD) simulations

The structure of WT intimin (PDB entry 4e1s; Fairman *et al.*, 2012) was placed in a 140-*n*-dodecylphosphocholine (DPC) micelle using *CHARMM-GUI* (Brooks *et al.*, 2009; Jo *et al.*, 2008). The DPC molecules were then mutated to DDM using a modified topology based on the CHARMM force field. Mutations for intimin-EE1 and the additional alanine insertions for intimin-EE2 were made *in silico* using *VMD* (Humphrey *et al.*, 1996) and were minimized using *NAMD* (Phillips *et al.*, 2005). An eight-residue epitope was docked to Fab/EE using *ClusPro*2 (Comeau *et al.*, 2004). Equilibration of this bound state over 10 ns was found to be stable and thus was used to model the placement of the Fab/EE on both intimin-EE1 and intimin-EE2. The systems were then each solvated with water and ionized with 250 m*M* NaCl. The resulting systems contained approximately 172 000 atoms each.

Restraints were applied for a short equilibration to ensure that the epitope-containing loop had stabilized in the Fab/EE binding site. The resulting systems were then run using unrestrained MD for 50 ns for intimin-EE1 and for 170 ns for intimin-EE2. All simulations were run using *NAMD*. The CHARMM36 (Best *et al.*, 2012) force field was used throughout, along with the TIP3P water model (Jorgensen *et al.*, 1983). A 2 fs time step was used for all bonded and short-range interactions, with long-range nonbonded electrostatics calculated every other time step using the particle mesh Ewald method (Darden *et al.*, 1993). A uniform temperature of 310 K and pressure of 101 325 Pa were maintained.

3. Results and discussion

3.1. Fab/EE rationale, cloning and biophysical properties

Previously, we reported the successful engineering of both EE-specific and hexahistidine peptide-specific scFvs for potential use as crystallization chaperones (Pai *et al.*, 2011; Kalyoncu *et al.*, 2014). Ultimately, the anti-EE scFv proved to have higher affinity and to be preferable over its anti-hexahistidine counterpart likely owing to its chemical diversity, its insensitivity to pH especially at near-physiological conditions and its greater compatibility with a variety of peptide-insert locations (terminal and internal). Here, we sought to convert our EE-specific scFv to the Fab format, which represents the majority of crystallization chaperone proteins in solved structures deposited in the PDB. Fab molecules also typically exhibit better biophysical properties such as enhanced thermal stability and well defined oligomeric states compared with the scFv platform (Röthlisberger *et al.*, 2005). We also postulated that the larger size of a Fab (∼50 *versus* ∼25 kDa) would provide additional epitopes to

Table 2
Biophysical characteristics of EE peptide-binding antibody fragments.

Parameter	scFv/EE†	Fab/EE
Expression level (mg per litre of culture)	2.1	2.4
Solubility (mg ml^{-1})	12.8	9.4
Melting temperature (°C)	47.2 ± 0.3	59.8 ± 0.1
Monomeric protein (%)	81	87

† As reported in Pai *et al.* (2011).

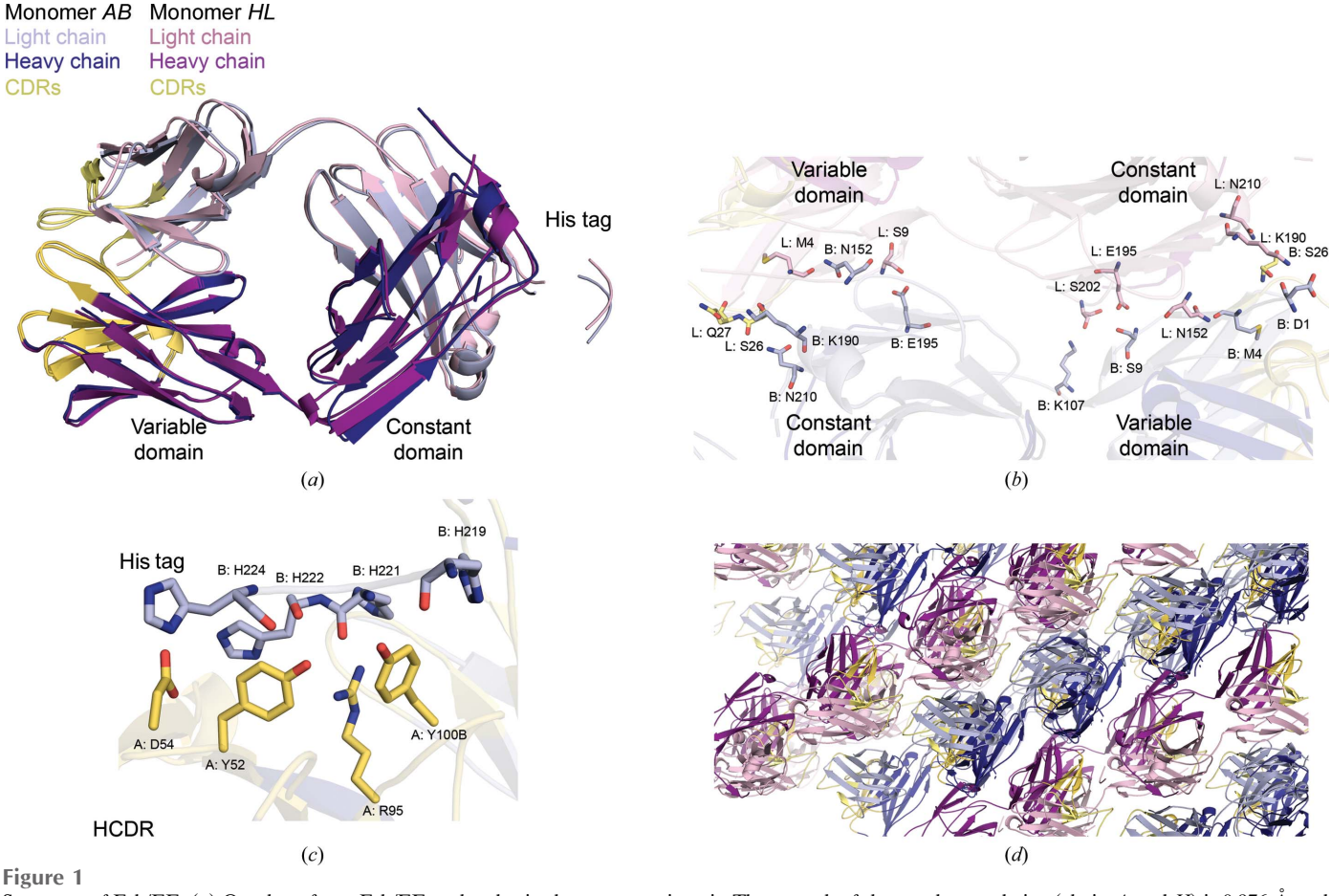


Figure 1
Structure of Fab/EE. (*a*) Overlay of two Fab/EE molecules in the asymmetric unit. The r.m.s.d. of the two heavy chains (chain *A* and *H*) is 0.976 Å and the r.m.s.d. of the two light chains (chain *B* and *L*) is 0.599 Å. (*b*) Fab/EE crystal contact ID2 (Table 3). The interface is 1109 Å^2 and includes nine residues from chain *L* and nine residues from chain *B* (both light chains; interacting residues are modeled as sticks). (*c*) Crystal contact ID5 (see Table 3). Interaction between the modelled decahistidine tag from chain *L* and CDR of chain *H* of a different molecule. Interacting residues are modeled as sticks. (*d*) Fab/EE lattice showing extended crystal contact areas and the lack of a channel that could accommodate a membrane protein. Colors are as in (*a*).

mediate crystal contacts, which would be especially advantageous for larger client membrane proteins.

Fab/EE was generated by subcloning the variable regions into the dicistronic plasmid pFab for bacterial expression (Levy *et al.*, 2001). Here, each variable region is appended with a cognate human constant domain such that both polypeptide chains (V_L–C_L and V_H–C_H1) are targeted to the periplasm for correct assembly and disulfide-bond formation. After osmotic shock and purification (Supplementary Fig. S1), similar levels of total and monomeric Fab/EE protein were recovered as for scFv/EE (2.4 *versus* 2.1 mg per litre of culture; 87 *versus* 81% monomeric, respectively; Table 2). Fab/EE thermal stability was significantly enhanced compared with scFv/EE (Table 2). The combination of high expression yields and enhanced stability, a biophysical characteristic correlated with crystallization success rates (Ericsson *et al.*, 2006), render Fab/EE more promising than the parent scFv/EE for use in large-scale co-crystallization trials.

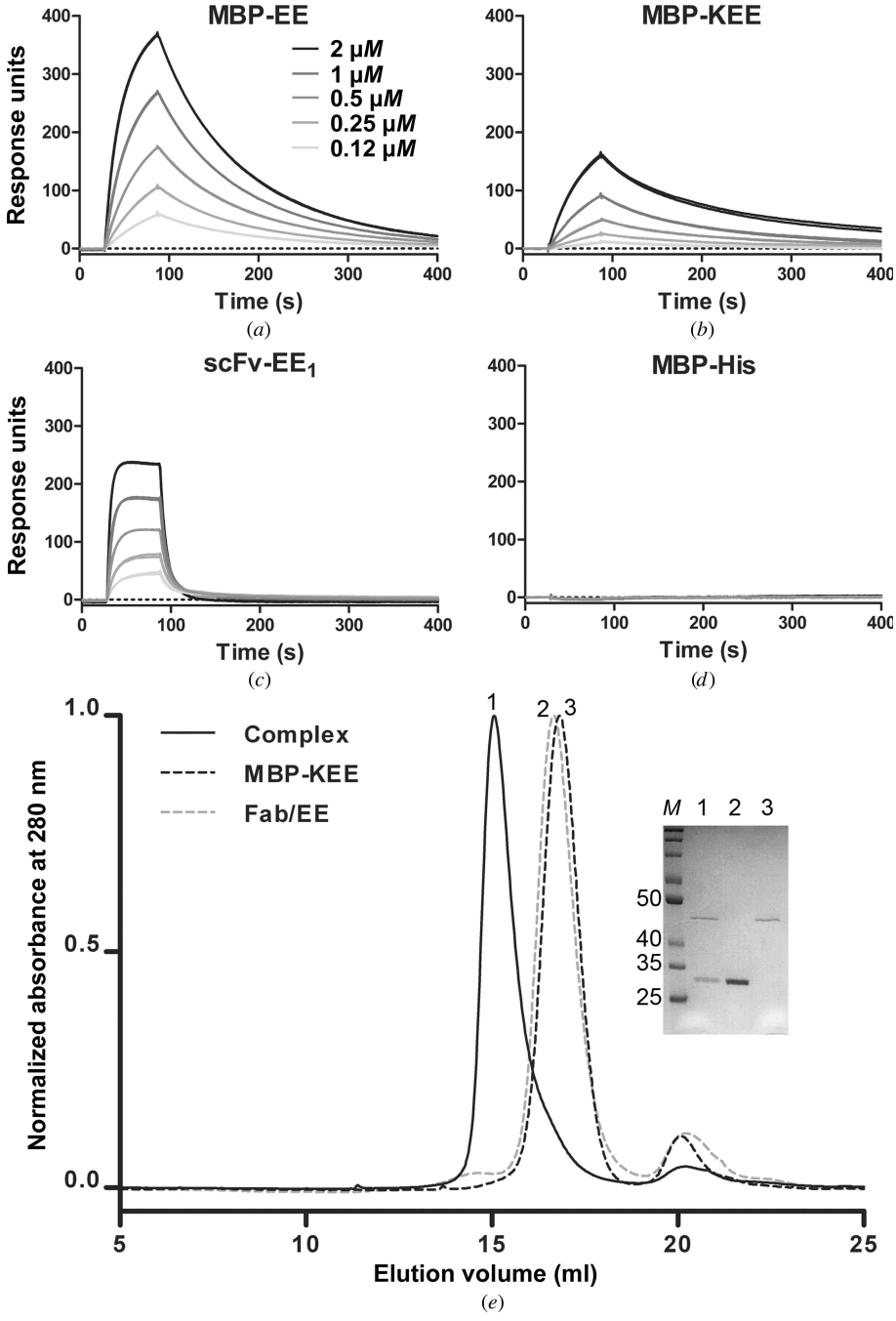


Figure 2
Fab/EE binding to EE-tagged soluble proteins. EE peptide-containing soluble proteins were injected in duplicate for each concentration tested and binding to immobilized Fab/EE was monitored by SPR. Both duplicates traces are shown and demonstrate Fab/EE binding to (*a*) MBP-EE, (*b*) MBP-KEE and (*c*) scFv-EE_1. (*d*) No binding was observed for MBP-His_6. (*e*) Equimolar amounts of MBP-KEE and Fab/EE were equilibrated at room temperature separately or together prior to separation by SEC. The sample containing both proteins eluted as a single peak with a shorter retention volume than either individual protein. Inset: fractions from each major peak were analyzed by reducing SDS–PAGE. Lane 1, peak fraction from the putative complex peak. Lane 2, peak fraction from the Fab/EE peak. Lane 3, peak fraction from the MBP-KEE peak. Lane *M* contains molecular-weight marker (labeled in kDa).

3.2. Fab/EE structural characterization

Crystals of Fab/EE belonged to space group *P*1 and the 2.0 Å resolution structure was solved by molecular replacement. Most residues were successfully modelled into the electron-density map. Exceptions include Ser128–Ser134 in chain *H*, the FLAG tag at the C-terminal end of the heavy chains (chains *H* and *A*), the linker residues between the C-terminal ends of the light chains (chains *L* and *B*) and the last four residues of the decahistidine-tag (Supplementary Table S1 and S2). The only Ramachandran outlier in the Fab/EE structure is His222, which is part of the decahistidine tag of chain *B* and fits well into the electron density. Two nearly identical molecules are present in the asymmetric unit. Chains *L* and *B* (light chains) superimpose with a root-mean-square deviation (r.m.s.d.) of 0.599 Å and chains *A* and *H* (heavy chains) with an r.m.s.d. of 0.976 Å using the secondary-structure matching (SSM) function within *Coot* (Krissinel & Henrick, 2007; Fig. 1*a*). Fab/EE chains *A* and *B* (monomer *AB*) and scFv/EE chains *A* and *B* (PDB entry 3nn8) are also nearly identical, with an r.m.s.d. of 0.708 Å between the two light chains and of 0.581 Å between the two heavy chains.

The lattice of Fab/EE demonstrates the variety of crystal contacts available

Table 3
Crystal lattice contact comparison for scFv/EE and Fab/EE.

Antibody fragment	*PISA* interface	Interface area (Å^2)	Contacts that form hydrogen bonds/salt bridges			
			V_H	V_L	C_H1	C_L
scFv/EE†	ID2	465	Lys57, Arg58, Ser65, Thr68	—	—	—
	ID3	456	Pro9, Lys19, Ser30, Ser32, Asp72, Lys73, Ser75, Thr77, Tyr79, Ser98	—	—	—
	ID4	393	Gln6, Asp11, Tyr91, Gln105, Thr108	Leu106, Lys107, Arg108, Gly110 (linker)	—	—
Fab/EE	ID2	1109	—	Asp1, Met4, Ser9, Ser26, Gln27, Lys107	—	Asn152, Lys190, Glu195, Ser202, Asn210
	ID3	517	Gln1, Ser28, Val71,	—	Glu150, Lys213	—
	ID4	345			Ser168, Asn216	
	ID5	325	Tyr52, Trp53, Asp54, Asp56, Arg95, Tyr100B	—	—	His219 (His tag), His221 (His tag), His222 (His tag), His224 (His tag)

† As reported in Kalyoncu *et al.* (2014).

to aid in the crystallization of client proteins (Figs. 1*b*, 1*c* and 1*d*; Table 3). Overall, the crystal contacts of Fab/EE had larger interface surface areas than those of scFv/EE, with the largest interface of Fab/EE calculated to be 1109 Å^2 (ID2 in Table 3; Fig. 1*b* and Supplementary Fig. S4). The only residue that participates in crystal contacts in both Fab/EE and scFv/EE is Lys107 (Fig. 1*b*), and the two interfaces are not similar. The Fab/EE interface contains the variable domain of chain *B* in one Fab/EE molecule interacting with the constant domain of chain *L* in the second Fab/EE molecule (Fig. 1*b*). The corresponding scFv/EE interface is between the variable domains of the light chain and the heavy chain. The next largest Fab/EE contact by surface area, ID3, has interactions between heavy chains of adjacent molecules (not shown), with less than half the interface area of ID2 (Table 3). Unexpectedly, a portion of the decahistidine tag on the C-terminus of the light chain of Fab/EE forms hydrogen-bonding and salt-bridge interactions (ID5) with the complementarity-determining regions (CDRs) from the heavy chain. This interaction is seen in both molecules in the asymmetric unit (Fig. 1*c* and Supplementary Fig. S5). Since the negative control proteins for complexation studies all have a histidine tag (see the SPR and SEC data in §§3.3 and 3.5 below), we surmise that this is an artifact of crystallization. Finally, the *P*1 lattice lacks solvent channels to accommodate a client EE-tagged protein (Fig. 1*d*) and the CDRs are being used in the crystal contacts. Upon binding to target protein, Fab/EE would likely utilize other residues that are available for forming crystal contacts, perhaps including the crystal contact areas seen in scFv/EEs (Kalyoncu *et al.*, 2014).

3.3. Fab/EE forms stable complexes with soluble EE-tagged client proteins

To assess the EE-binding properties of Fab/EE, we first examined interactions with soluble proteins presenting the EE peptide by SPR. A binding affinity in the nanomolar range was measured for Fab/EE binding to MBP containing a C-terminal (MBP-EE) or an internal (MBP-KEE) EE peptide, as well as to an scFv with the EE peptide inserted into the flexible linker region (scFv-EE$_1$) (Table 4, Figs. 2*a*, 2*b* and 2*c*). No significant affinity loss was observed upon conversion of the scFv format to a Fab [767 n*M* for scFv/EE binding to MBP-EE (Pai *et al.*, 2011) and 308 n*M* for Fab/EE binding to MBP-EE], as has been observed previously (Adams & Schier, 1999; Maynard *et al.*, 2002; Krebs *et al.*, 2001). No binding was observed to the negative control MBP-His$_6$ (Fig. 2*d*).

Since SPR employs surface-immobilization strategies that may not accurately reflect binding in solution or in a crystallization drop, we assessed complex stability by SEC. Notably, SEC is often used to isolate protein–protein complexes prior to co-crystallization trials (Lim *et al.*, 2011; Roosild *et al.*, 2006). MBP-KEE was selected among the aforementioned client proteins for solution complexation with Fab/EE. A clear shift in elution volume corresponding to a higher molecular mass was observed for the complex, and complexation was confirmed by SDS–PAGE (Fig. 2*e*).

3.4. EE-peptide insertion into the extracellular loop of membrane proteins

Next, we examined complex formation with the model membrane proteins α-helical A$_2$aR and the β-barrel intimin. To determine candidate locations for EE-peptide insertion into A$_2$aR, the repertoire of previously solved GPCR structures was analyzed, revealing that most previously solved GPCR structures involved modification of the flexible ICL3 (Rosenbaum *et al.*, 2009), either by truncation or through the use of covalent (T4L or BRIL; Rosenbaum *et al.*, 2007) or noncovalent (Fab; Rasmussen *et al.*, 2007) crystallization chaperones targeting ICL3 identified by hybridoma technology (Day *et al.*, 2007). Specifically, ten of the 12 residues that form the Fab epitope in the complex structure with the β_2 adrenergic receptor (Rasmussen *et al.*, 2007) are contiguous in primary sequence, suggesting that the EE peptide inserted into the homologous location in A$_2$aR would be accessible by our Fab/EE. Thus, the EE peptide flanked by flexible linkers (sequence GS-EYMPME-GS) was inserted into the N-terminal segment of ICL3 of an A$_2$aR-green fluorescence protein construct after Lys209 without mutation or truncation of the wild-type residues (Supplementary Fig. S2*a*) to generate A$_2$aR-GFP-EE. A$_2$aR-GFP and A$_2$aR-GFP-EE

Table 4
Characterization of EE-tagged protein binding kinetics of Fab/EE by SPR.

	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d (nM)
MBP-EE	$3.39 \pm 1.23 \times 10^{4}$	$9.22 \pm 0.15 \times 10^{-3}$	308 ± 117
MBP-KEE	$1.25 \pm 0.30 \times 10^{4}$	$7.88 \pm 0.91 \times 10^{-3}$	612 ± 95
scFv-EE_1	$2.95 \pm 1.47 \times 10^{5}$	$5.56 \pm 1.66 \times 10^{-2}$	224 ± 160
A_2aR-GFP-EE	$2.13 \pm 1.60 \times 10^{5}$	$3.92 \pm 1.96 \times 10^{-3}$	32 ± 24

were expressed and purified from *S. cerevisiae*. Silver-staining SDS–PAGE reveals predominantly monomeric A_2aR-GFP-EE and Western blot shows EE-peptide incorporation into only A_2aR-GFP-EE and not wild-type A_2aR-GFP (Supplementary Fig. S2*b*).

The representative β-barrel membrane protein, *E. coli* intimin, was selected because of its high recombinant expression yield but its relative difficulty in crystallization (Fairman *et al.*, 2012). Intimin was engineered with an internal EE peptide (residues 315–320 of the L4 loop region were mutated to EYMPME; intimin-EE1) as reported previously (Kalyoncu *et al.*, 2014). Analogous to A_2aR, a second construct possessing an insertion of two alanine residues on either side of the EE peptide was generated (Supplementary Fig. S3*a*; intimin-EE2), which we postulated would increase the accessibility of the EE peptide for better Fab/EE binding. Both intimin-EE1 and intimin-EE2 were expressed and purified (Supplementary Fig. S3*b*) with similar protein yields to WT intimin (Fairman *et al.*, 2012).

3.5. Fab/EE forms complexes with EE-tagged membrane proteins

The accessibility and specific binding of Fab/EE to the EE peptide as presented by A_2aR-GFP was limited to ELISA (Fig. 3*a*) and SPR (Figs. 3*b*, 3*c* and 3*d*) owing to the low yields of purified protein. For both experiments Fab/EE binding was only detected for EE-tagged proteins. The calculated K_d from SPR for A_2aR-GFP-EE is lower than the K_d for the soluble

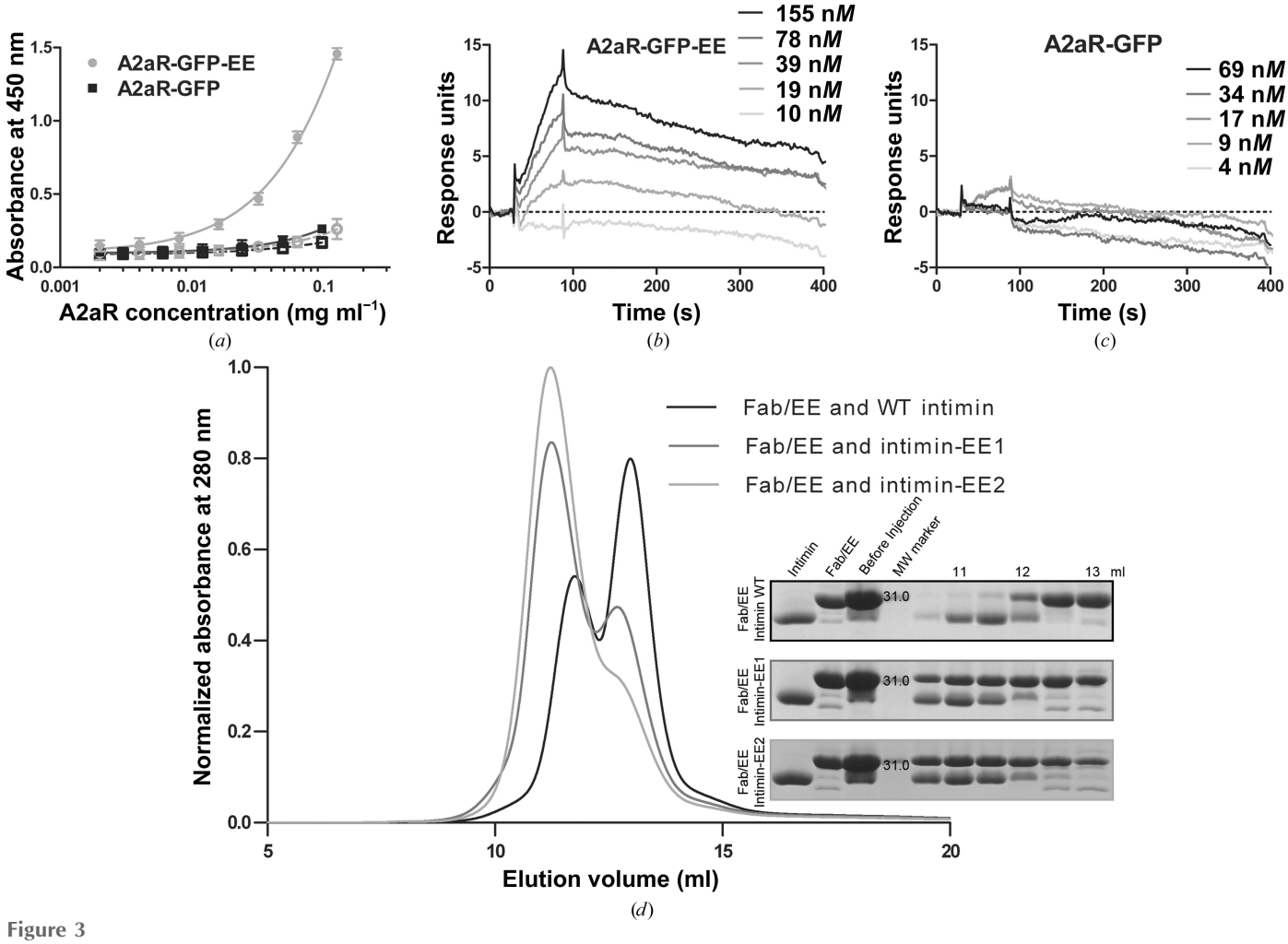


Figure 3
Fab/EE binding to EE-tagged membrane proteins. (*a*) ELISA analysis of Fab/EE binding to A_2aR-GFP. Purified A_2aR-GFP proteins with or without the EE-peptide insertion were added to ELISA wells coated with Fab/EE or blocked control wells. Captured protein was detected using an anti-GFP antibody. (*b*, *c*) SPR analysis of binding. GPCRs were injected in duplicate for each concentration tested and binding to immobilized Fab/EE was monitored. Average traces are shown and demonstrate concentration-dependent binding to A_2aR-GFP-EE. No binding was observed for A_2aR-GFP lacking the EE peptide. (*d*) Elution profile of purified Fab/EE incubated with WT intimin, intimin-EE1 or intimin-EE2. Reducing SDS–PAGE analysis (inset) of fractions selected from 10.5 to 13 ml elution volumes shows coelution of Fab/EE–intimin-EE1 and Fab/EE–intimin-EE2.

proteins MBP-EE, MBP-KEE and scFv-EE$_1$ (Table 4), perhaps because of the presence of detergent.

The ability of Fab/EE to form a solution complex with an EE-tagged membrane protein was tested by SEC with the intimin constructs (Fig. 3*d*). The Fab/EE–WT intimin elution trace has two peaks, the first representing WT intimin and the second Fab/EE, as confirmed by SDS–PAGE (top of the inset in Fig. 3*d*). By contrast, when intimin-EE1 or intimin-EE2 is mixed with Fab/EE, the species coelute. The first peak in the SEC trace has a shorter retention time indicative of a higher molecular-mass complex expected for Fab/EE–intimin-EE complexes, concomitant with a reduction in the peak corresponding to Fab/EE alone (middle and bottom of the inset in Fig. 3*d*).

3.6. MD simulations of Fab/EE–intimin-EE complexes

We postulated that upon binding of Fab/EE to intimin-EE1 and intimin-EE2, loop L4 residues would be immobilized and a stable complex for crystallization would be generated. Over 4000 crystallization trials of Fab/EE and intimin-EE variants have been set up (∼35 × 96 conditions from commercial screens and 29 × 24 conditions from homemade screens), with varying temperatures, concentrations and methods [vapour

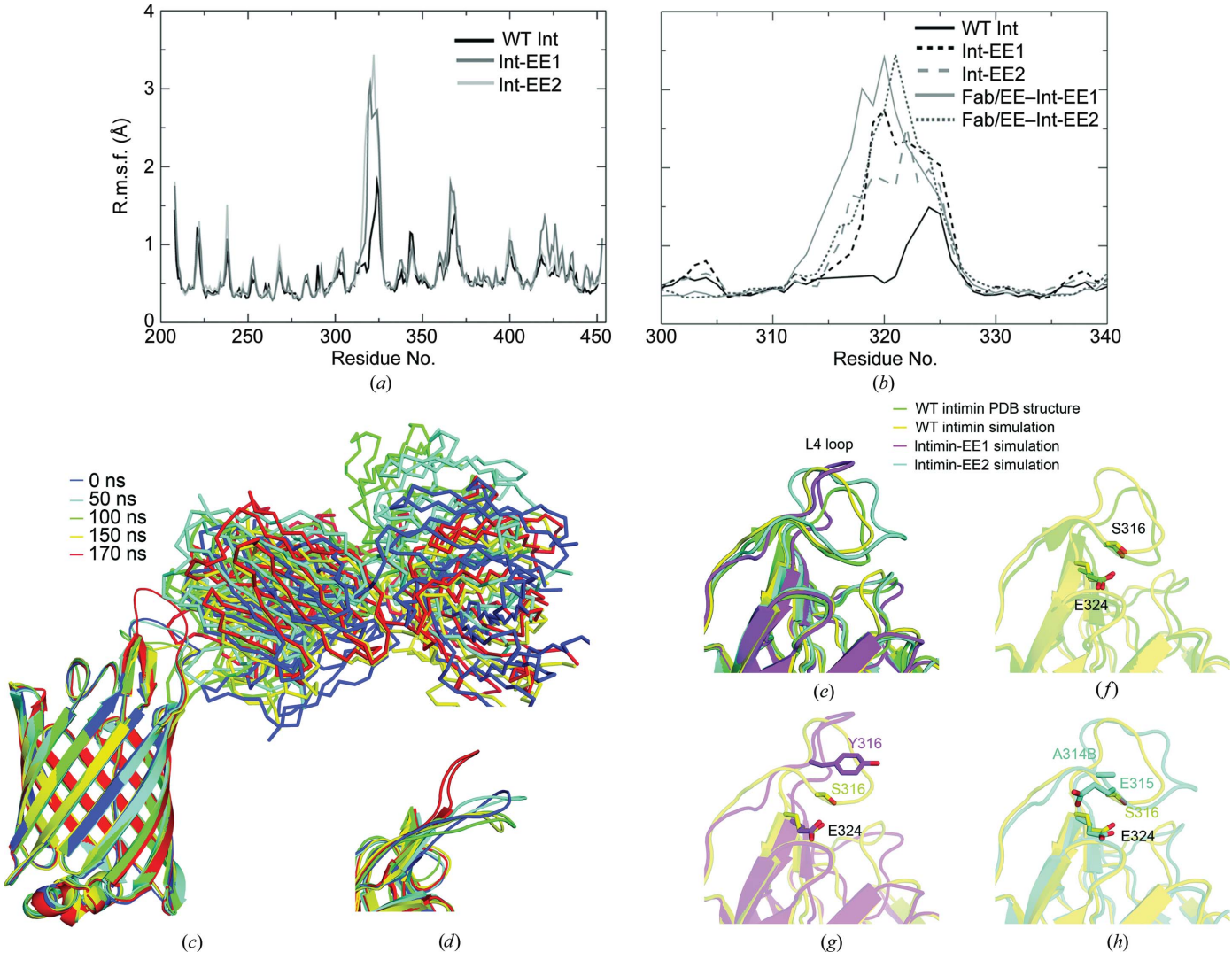


Figure 4
Molecular-dynamics analysis of intimin-EE1 and intimin-EE2 with and without Fab/EE. (*a*) R.m.s.f. of residues 200–450 of WT intimin, intimin-EE1 and intimin-EE2 over 50 ns simulation. (*b*) R.m.s.f. of residues 300–340 for WT intimin and for intimin-EE1 and intimin-EE2 alone and in modelled complexes with Fab/EE over 50 ns simulation. (*c*) Proposed Fab/EE–intimin-EE2 complex structure at different time points during the simulation, colored from blue to red. Intimin-EE2 is shown as a cartoon and Fab/EE as a ribbon. (*d*) An enlarged view of the intimin-EE2 loop containing the EE peptide at the same time points as in (*c*) with Fab/EE omitted for clarity. (*e*) L4 loops of intimin structures: WT intimin (PDB entry 4e1s) and WT intimin, intimin-EE1 and intimin-EE2 after 50 ns simulation. (*f*) The L4 loop of WT intimin from the PDB structure and WT intimin after 50 ns simulation. Ser316 hydrogen bonds to Glu324 in PDB entry 4e1s and the interaction is unchanged after 50 ns simulation. (*g*) WT intimin and intimin-EE1 after 50 ns equilibration showing L4 loop residue 316 and Glu324 for each structure. (*h*) WT intimin and intimin-EE2 after 50 ns equilibration showing L4 loop residues Ser316 and Glu324 of WT intimin, Ala314B and Glu315 of intimin-EE2.

diffusion with detergent (DDM), vapour diffusion with bicelle (CHAPSO/DMPC) or lipidic cubic phase (monoolein), data not shown]. Both direct mixing of 1:1 Fab/EE–intimin-EE variant immediately prior to setting up trays and using the complex isolated from SEC as above were used in co-crystallization trials. Unfortunately, none yet has resulted in diffracting crystals of the Fab/EE–intimin-EE complex; any initially promising leads obtained to date have ultimately been crystals of Fab/EE alone (data not shown).

To gain insight into why the co-crystallization trials have not been successful in spite of the demonstrated favorable solution properties, we turned to molecular dynamics. WT intimin, intimin-EE1 and intimin-EE2 were first modelled and allowed to equilibrate over 50 ns (Fig. 4*a*). In comparison to WT intimin, intimin-EE1 and intimin-EE2 exhibit increased flexibility in residues 315–320, where the WT residues were mutated to the EE tag. Notably, the insertion of two alanine residues on each side of the EE tag in intimin-EE2 did not seem to cause a further increase in flexibility compared with intimin-EE1. After modeling the Fab/EE–intimin-EE interaction, the Fab/EE–intimin-EE1 complex was equilibrated for 50 ns and the Fab/EE–intimin-EE2 complex was equilibrated for 170 ns. The r.m.s.f. (root-mean-squared fluctuation) of residues 300–340 after 50 ns with and without Fab/EE bound (Fig. 4*b*) indicates that Fab/EE binding to both intimin-EE1 and intimin-EE2 slightly increases the flexibility of residues 312–325. Over the 170 ns of simulation for the Fab/EE–intimin-EE2 complex, the position of Fab/EE in relation to intimin-EE2 is dynamic (Figs. 4*c* and 4*d*). Such flexibility is likely to be a result of mutations in the L4 loop in the membrane protein (Fig. 4*e*). In particular, Ser316 of WT intimin forms hydrogen-bonding interactions in both WT intimin and the structure after 50 ns of MD simulations (Fig. 4*f*). This interaction was broken when Ser316 was mutated to tyrosine within EYMPME of intimin-EE1, resulting in drastic changes within the loop (Fig. 4*g*). The corresponding residue in intimin-EE2 is Ala314B (see the numbering in Supplementary Fig. S2*a*) and after 50 ns simulation the loop resembles that of WT intimin but with Glu315 of intimin-EE2 in a similar position as the WT intimin residue Ser316 (Fig. 4*h*).

4. Conclusions: implications for chaperone-mediated crystallization of MPs

Our engineered Fab/EE exhibits numerous favorable characteristics for use as a crystallization chaperone for difficult targets. Introduction of the EE epitope within membrane proteins is straightforward by SDM, leading to readily detected high-affinity solution complexes. However, even when placement of the EE epitope does not interfere with protein expression and purification, an unintended consequence may be the removal of native contacts and thus an increase in conformational heterogeneity that is detrimental to crystallographic efforts. Such conformational changes can be challenging to predict, even for test proteins whose structures are known, and would be especially enigmatic in the case of target proteins of unknown structure. If an inflexible region of the target protein is not known, a reduction in loop flexibility is likely to be achievable by shortening the EE epitope-containing loop to a minimum number of residues that can still be complexed with Fab/EE. This approach, on its own and in combination with others including SER and stabilizing ligands, is under way in our laboratory.

Acknowledgements

We thank Dr Anne S. Robinson (Tulane University) for the pITy-A_2aR-GFP plasmid, Dr George Georgiou (University of Texas at Austin) for the pFab plasmid and Dr Susan Buchanan for the intimin plasmid. This work was supported by grants from the NIH (DK091357 to RLL, GM095638 to JAM and K22-AI100927 to JCG), NSF (0845445 to RLL), the Welch Foundation (No. F-1767 to JAM), the Department of Education Graduate Assistance in Areas of National Need (GAANN, P200A090307 to JLJ), the Petit Undergraduate Scholars program at Georgia Tech (DPH and IAM) and a Georgia Tech Senior Biophysics Training award (JLJ and SK). The use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. Simulations were carried out using Stampede at TACC through the Extreme Science and Engineering Discovery Environment (XSEDE).

References

Adams, G. P. & Schier, R. (1999). *J. Immunol. Methods*, **231**, 249–260.
Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979). *Biochemistry*, **18**, 5698–5703.
Banatao, D. R., Cascio, D., Crowley, C. S., Fleissner, M. R., Tienson, H. L. & Yeates, T. O. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 16230–16235.
Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M. & MacKerell, A. D. Jr (2012). *J. Chem. Theory Comput.* **8**, 3257–3273.
Bourhis, E. *et al.* (2011). *Structure*, **19**, 1433–1442.
Brooks, B. R. *et al.* (2009). *J. Comput. Chem.* **30**, 1545–1614.
Caffrey, M. (2003). *J. Struct. Biol.* **142**, 108–132.
Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Kuhn, P., Weis, W. I., Kobilka, B. K. & Stevens, R. C. (2007). *Science*, **318**, 1258–1265.
Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004). *Bioinformatics*, **20**, 45–50.
Connelly, P. R. (1994). *Curr. Opin. Biotechnol.* **5**, 381–388.
Costa, G. L. & Weiner, M. P. (2006). *Cold Spring Harb. Protoc.* **2006**, doi:10.1101/pdb.prot4144.
Darden, T., York, D. & Pedersen, L. (1993). *J. Chem. Phys.* **98**, 10089–10092.
Day, P. W., Rasmussen, S. G. F., Parnot, C., Fung, J. J., Masood, A., Kobilka, T. S., Yao, X.-J., Choi, H.-J., Weis, W. I., Rohrer, D. K. & Kobilka, B. K. (2007). *Nature Methods*, **4**, 927–929.
Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. (2006). *Anal. Biochem.* **357**, 289–298.
Fairman, J. W., Dautin, N., Wojtowicz, D., Liu, W., Noinaj, N., Barnard, T. J., Udho, E., Przytycka, T. M., Cherezov, V. & Buchanan, S. K. (2012). *Structure*, **20**, 1233–1243.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.

Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M. B., Hovius, R., Graff, A., Stahlberg, H., Tomizaki, T., Desmyter, A., Moreau, C., Li, X.-D., Poitevin, F., Vogel, H. & Nury, H. (2014). *Nature (London)*, **512**, 276–281.

Howell, S. C., Mittal, R., Huang, L., Travis, B., Breyer, R. M. & Sanders, C. R. (2010). *Biochemistry*, **49**, 9572–9583.

Huber, T., Steiner, D., Röthlisberger, D. & Plückthun, A. (2007). *J. Struct. Biol.* **159**, 206–221.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Hunte, C., Koepke, J., Lange, C., Rossmanith, T. & Michel, H. (2000). *Structure*, **8**, 669–684.

Jo, S., Kim, T., Iyer, V. G. & Im, W. (2008). *J. Comput. Chem.* **29**, 1859–1865.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). *J. Chem. Phys.* **79**, 926–935.

Kalyoncu, S., Hyun, J., Pai, J. C., Johnson, J. L., Entzminger, K., Jain, A., Heaner, D. P. Jr, Morales, I. A., Truskett, T. M., Maynard, J. A. & Lieberman, R. L. (2014). *Proteins*, **82**, 1884–1895.

Keenan, R. J., Siehl, D. L., Gorton, R. & Castle, L. A. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8887–8892.

Koldobskaya, Y., Duguid, E. M., Shechner, D. M., Suslov, N. B., Ye, J., Sidhu, S. S., Bartel, D. P., Koide, S., Kossiakoff, A. A. & Piccirilli, J. A. (2011). *Nature Struct. Mol. Biol.* **18**, 100–106.

Kovari, L. C., Momany, C. & Rossmann, M. G. (1995). *Structure*, **3**, 1291–1293.

Krebs, B. *et al.* (2001). *J. Immunol. Methods*, **254**, 67–84.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Laganowsky, A., Zhao, M., Soriaga, A. B., Sawaya, M. R., Cascio, D. & Yeates, T. O. (2011). *Protein Sci.* **20**, 1876–1890.

Levy, R., Weiss, R., Chen, G., Iverson, B. L. & Georgiou, G. (2001). *Protein Expr. Purif.* **23**, 338–347.

Lieberman, R. L., Culver, J. A., Entzminger, K. C., Pai, J. C. & Maynard, J. A. (2011). *Methods*, **55**, 293–302.

Lim, H.-H., Fang, Y. & Williams, C. (2011). *PLoS One*, **6**, e24653.

Link, A. J. & LaBaer, J. (2011). *Cold Spring Harb. Protoc.* **2011**, doi:10.1101/pdb.prot5651.

Liu, W., Chun, E., Thompson, A. A., Chubukov, P., Xu, F., Katritch, V., Han, G. W., Roth, C. B., Heitman, L. H., IJzerman, A. P., Cherezov, V. & Stevens, R. C. (2012). *Science*, **337**, 232–236.

Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.

Maynard, J., Adams, E. J., Krogsgaard, M., Petersson, K., Liu, C. W. & Garcia, K. C. (2005). *J. Immunol. Methods*, **306**, 51–67.

Maynard, J. A., Maassen, C. B. M., Leppla, S. H., Brasky, K., Patterson, J. L., Iverson, B. L. & Georgiou, G. (2002). *Nature Biotechnol.* **20**, 597–601.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.

O'Malley, M. A., Lazarova, T., Britton, Z. T. & Robinson, A. S. (2007). *J. Struct. Biol.* **159**, 166–178.

Ostermeier, C., Iwata, S., Ludwig, B. & Michel, H. (1995). *Nature Struct. Biol.* **2**, 842–846.

Pai, J. C., Culver, J. A., Drury, J. E., Motani, R. S., Lieberman, R. L. & Maynard, J. A. (2011). *Protein Eng. Des. Sel.* **24**, 419–428.

Pédelacq, J.-D., Piltch, E., Liong, E. C., Berendzen, J., Kim, C.-Y., Rho, B.-S., Park, M. S., Terwilliger, T. C. & Waldo, G. S. (2002). *Nature Biotechnol.* **20**, 927–932.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L. & Schulten, K. (2005). *J. Comput. Chem.* **26**, 1781–1802.

Piccirilli, J. A. & Koldobskaya, Y. (2011). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2918–2928.

Privé, G. G. (2009). *Curr. Opin. Struct. Biol.* **19**, 379–385.

Rasmussen, S. G. F. *et al.* (2011). *Nature (London)*, **469**, 175–180.

Rasmussen, S. G. F., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I. & Kobilka, B. K. (2007). *Nature (London)*, **450**, 383–387.

Roosild, T. P., Castronovo, S. & Choe, S. (2006). *Acta Cryst.* F**62**, 835–839.

Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Yao, X.-J., Weis, W. I., Stevens, R. C. & Kobilka, B. K. (2007). *Science*, **318**, 1266–1273.

Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. (2009). *Nature (London)*, **459**, 356–363.

Röthlisberger, D., Honegger, A. & Plückthun, A. (2005). *J. Mol. Biol.* **347**, 773–789.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd ed. New York: Cold Spring Harbor Laboratory Press.

Sennhauser, G. & Grütter, M. G. (2008). *Structure*, **16**, 1443–1453.

Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.

Stura, E. A., Graille, M., Housden, N. G. & Gore, M. G. (2002). *Acta Cryst.* D**58**, 1744–1748.

Tereshko, V., Uysal, S., Koide, A., Margalef, K., Koide, S. & Kossiakoff, A. A. (2008). *Protein Sci.* **17**, 1175–1187.

Thompson, A. A., Liu, W., Chun, E., Katritch, V., Wu, H., Vardy, E., Huang, X.-P., Trapella, C., Guerrini, R., Calo, G., Roth, B. L., Cherezov, V. & Stevens, R. C. (2012). *Nature (London)*, **485**, 395–399.

Uysal, S., Vásquez, V., Tereshko, V., Esaki, K., Fellouse, F. A., Sidhu, S. S., Koide, S., Perozo, E. & Kossiakoff, A. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 6644–6649.

Vedadi, M. *et al.* (2006). *Proc. Natl Acad. Sci. USA*, **103**, 15835–15840.

Wernimont, A. & Edwards, A. (2009). *PLoS One*, **4**, e5094.

Wu, B., Chien, E. Y. T., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V. & Stevens, R. C. (2010). *Science*, **330**, 1066–1071.

Yang, J. K., Park, M. S., Waldo, G. S. & Suh, S. W. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 455–460.

Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. (2001). *Nature (London)*, **414**, 43–48.

Zou, Y., Weis, W. I. & Kobilka, B. K. (2012). *PLoS One*, **7**, e46039.